DOCUMENT RESUME

ED 060 820 HE 002 965

AUTHOR	Kennedy, Robert W.
TITLE	The Relationship of Selected Student Characteristics to Components of Teacher/Course Evaluations Among Freshman English Students at Kent State University.
INSTITUTION	Case Western Reserve Univ., Cleveland, Ohio.
PUB DATE	Apr 72
NOTE	31p.; Paper presented at the annual meeting of the American Education Research Association, Chicago, April, 1972
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*College Students; *College Teachers; *Evaluation; *Faculty; *Higher Education; Performance Specifications

ABSTRACT

The present study was conducted to determine the relationship between student ratings on the components of a teacher/course evaluation instrument and their scores on selected Omnibus Personality Inventory Subscales, American College Test scores, "expected grade," "actual grade," "expected-actual" grade differential in the course, gradepoint average, and the variables of sex and college membership. The research was completed using both standardized and nonstandardized instruments administered to freshmen students enrolled in a required English course during the 1970 fall quarter at Kent State University. The results are reported in a series of 37 tables. Suggestions for further, broader research in the area are made to determine what criteria variables students use to evaluate above-average teachers. (Author/HS)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 060820

The Relationship of Selected Student Characteristics to Components of Teacher/Course Evaluations Among Freshman English Students at Kent State University

HE
N-NPD

W. Robert Kennedy
Case Western Reserve University

The objective of this study was to determine the relationships between student ratings on the components of a teacher/course evaluation instrument and their scores on selected Omnibus Personality Inventory subscales, American College Test scores, "expected grade," "actual grade," "expected-actual grade" differential in the course, grade point average, and the variables sex and college membership. This research was completed using both standardized and non-standardized instruments administered to freshman students enrolled in English 160, a required course, Fall Quarter, 1970 at Kent State University.

Description of the Sample

Course Selection. The optimum situation for a study of this type would have been to have had a single teacher evaluated by a large number of students. Unfortunately, no required large section courses were being taught at that time in which a teacher was exclusively responsible for all aspects of the course. English 160, however, although taught in small classes (average twenty-five students per section) by one instructor, was available. In general the instructors were Teaching Fellows nearing the completion of their doctoral program with a background in teaching (college level and other). In addition, since each had two sections involved in the study, the possibility of having fifty student evaluations (at maximum) per teacher made this course a meaningful selection. Considering the relatively common college background and experience of the students in this course, the usefulness of the data collected across fifteen teachers who were teaching the comparable content was seen as potentially much more valuable than that available in samples

HE 002 965

This paper was presented at the meeting of the American Education Research Association, Chicago, April, 1972

derived from different course offerings.

Teacher Sample. The sample of students used was contingent upon individual teachers volunteering to participate in the study. Teachers who were to be teaching two sections of English 160 were asked to volunteer. Fifteen of a possible one hundred eighty did so and all were included in the sample. Demographically, this group consisted of eight males and seven females, two with B.A. or B.S. degrees, eleven in post-Master's course work and two with the Ph.D. degree. The range for college teaching experience was zero to six years with the median being two years.

Student Sample. The number of freshman students enrolled in English 160 numbered nearly 750 (15 teachers x 50 students), however, the actual number included in the final sample was 549. Among the reasons for this decline in the sample size were the failure of students to volunteer for participation, absenteeism on testing dates, unusable forms, and the usual first-quarter freshman attrition rate.

Demographically the final sample of students included 222 males and 327 females. Of this group, 132 gave their college membership as Arts and Sciences; Fine and Professional Arts - 108; Business - 59; Education - 155; Nursing - 28; Health, Physical Education and Recreation - 6. The final grade distribution for English 160 (based on information supplied by thirteen teachers) included 50 A's; 165 B's; 220 C's; 32 D's and no failures.

Instruments

Teacher/Course Evaluation Instrument. The evaluative rather than the behavioral approach was used as the basis for the construction of this instrument with the item format reflecting "how adequately" the students felt the teacher had performed rather than whether or not a specific teacher activity had occurred.

The final evaluation instrument consisting of 45 items (see Appendix A) was the result of pretesting 78 items using a sample of 80 students enrolled in English

161 during Summer Session II, 1970. One half of the sample was asked to evaluate the items for their applicability in rating teachers in this course while the remainder were requested to rate their teachers' performance that session. Unfortunately, this unexpectedly small sample made factor analysis for scale development untenable, and as an alternative, item variances using both ratings were compared. Using as a basis small variances on the former and large variances on the latter, items were selected for inclusion in the revised instrument.

The total sample from the Fall Quarter testing included 549 teacher/course evaluation forms. These results were then analyzed using the principle components analysis technique with 1.00 values inserted in the diagonal. The matrix of intercorrelations included 36 items since the 9 items on the teacher personality subscale were deleted for this analysis. Eigen-values greater than 1.00 were used to determine the original factor stucture after which varimax rotations were used to determine the best alignment of the items. Several rotations were made, however, the four factor solution presented the best scale definition on the first three factors. The fourth factor, as seen in Table 4, was deleted from the final definition of scales since it was not only difficult to interpret, but, in additon, was probably the result of the high inter-correlations of items 24 and 28 rather than reflecting any substantive factor. The .3500 cutoff, rather than the traditional .3000 value, was used to determine the inclusion of items on each scale since the logical consistency of the scales was increased when items below .3500 were deleted. The number of items per scale was not appreciatively altered by this approach. For example, five items were deleted from Subscale I, none from Subscale II, and only four from Subscale III using this approach.

The alignment of items and their factor loadings for each of the three components are presented in Tables 1 through 3. The first subscale, "Instructional Methods," includes fifteen items which reflect various aspects of the instructional methods or procedures used in the classroom. Only one item, number four, "made

TABLE 1

Listing of Items and Factor Loading Exceeding .3500 for Factor I - Instructional Methods

Loading	Item Number	Items
.7031	31	-was a monotonous and dull speaker.
.7008	6	-presented the material in an interesting manner.
.6920	15	-used a style of lecturing which was dull and uninteresting.
.6338	30	-presented the material so that it was intellectually challenging to the student.
.5939	1	-was able to effectively relate the course materials to the broader field of knowledge.
.5891	13	-was concerned about stimulating students' curiosity in the subject.
.5759	33	-was able to stimulate interesting class discussions.
.5549	5	-capably related information from other fields to the course material.
.5485	12	-showed an engaging enthusiasm for the subject.
.5040	18	-effectively used mimicry, anedotes, and/or other general hamminess to enliven the class period.
.4868	44	-How much do you feel you have learned from this teacher?
.3976	45	-How much do you think you would like the instructor in this course as a personal friend?
.3848	4	-made valid interpretations of the readings (e.g., poems, novels)
.3808	34	-displayed only a text related knowledge.
.3662	16	-effectively used a variety of instructional methods which were appropriate to the course material.

TABLE 2

Listing of Items and Factor Loadings Exceeding .3500 for Factor II - Interpersonal Relationships With Students

Loading	Item Number	Items
.6572	29	-was threatening and caused students to be afraid of speaking in class.
.6569	32	-often made individual students feel uncomfortable or embarrassed in class.
.5989	27	-was sometimes unfair in the grading of students' work.
.5829	26	-too often forced his (her) opinions on the class.
.5333	8	-presented the subject matter in too complex a manner.
.5144	22	-was too inflexible concerning his (her) right to control the in-class discussions and activities.
.4866	45	-How much do you think you would like the instructor in this course as a personal friend?
.4664	44	-How much do you feel you have learned from this teacher?
.4635	11	-by his (her) actions, seemed to view teaching as a chore or routine activity.
.4535	20	-relied too heavily on student performance (e.g., taking, answering questions, etc.) in class as the primary basis of grading.
.3928	7	-was able to effectively communicate the rules of good writing.
.3546	14	-often used class time with the discussion of irrelevent or meaningless topics.

TABLE 3

Listing of Items and Factor Loadings Exceeding .3500 for Factor III - Content Competency

Loading	Item Number	Items
.6203	3	-showed a good working knowledge of the rules of grammar.
.5944	7	-was able to effectively communicate the rules of good writing.
.5917	2	-expressed himself (herself) clearly when writing.
.5831	21	-was generally well prepared for class.
.5663	9	-was able to effectively synthesize, integrate, and summarize the subject matter.
.5624	19	-was able to communicate clearly the directions for the individual assignments.
.4776	17	-listened attentively to students' questions, comments, and remarks during class.
.4517	4	-made valid interpretations of the readings (e.g., poems, novels)
.4426	16	-effectively used a variety of instructional methods which were appropriate to the course material.
.4247	25	-was available for students to talk to when not in class.
.4143	1	-was able to effectively relate the course materials to the broader field of knowledge.
.3564	30	-presented the material so that it was intellectually challenging to the student.

TABLE 4

Listing of Item and Factor Loadings Exceeding .3500 for Factor IV

Loading	Item Number	Items
.6290	28	-was punctual about dismissing class.
.6209	24	-was punctual about meeting class.
.4114	34	-displayed only a text related knowledge.

TABLE 5

Listing of Adjective Pairs Used For Teacher Personality - Subscale IV

Item No.	Adjective Pair
35	Uneasy and Hesitant.....Calm and relaxed
36	Lacked sense of humor..Displayed sense of humor
37	Sincere.....Insincere
38	Self-confident.....Uncertain
39	Cold and impersonal.....Warm and friendly
40	Flexible.....Rigid
41	Threatening.....Non-threatening
42	Formal.....Informal
43	Unsociable.....Sociable

valid interpretations of the reading (e.g. poems, novels)," could be considered unique to classes in English as compared to other course offerings. Table 2 includes those items defining the "Interpersonal Relationships with Students" subscale and it also contains twelve items of which only one, item number seven, "was able to effectively communicate the rules of good writing," could be considered as rather unique to English classrooms. The third subscale, "Content Competency," includes the skills and abilities needed by teachers in this course. The transferability of the third scale to other content areas would probably be limited particularly when item three "showed a good working knowledge of the rules of grammer;" seven, "was able to effectively communicate the rules of good writing;" two, "expressed himself (herself) clearly when writing;" and four, "made valid interpretations of the readings (e.g. poems, novels)" are considered.

Table 5 is a presentation of the nine items on the "Teacher Personality" subscale. Because of the semantic differential rather than Likert format, these items were excluded from the component analysis and resulting scale construction.

The description of each of the three factors according to number of items, range of loadings, percent of variance accounted for in total factor space, and percent of total variance accounted for is presented in Table 6. Subscale I,

TABLE 6

Description of the Three Varimax Rotated Factors Determined for the Teacher/Course Evaluation Instrument

Names of Factors	Number of Items with Factor Loadings .3500	Range of Loadings	Percent of Variance Accounted for in the Common Factor Space of the Three Factors	Percent of the Total Variance Accounted For
I. Instructional Methods	15	.37 - .79	41.78	16.78
II. Interpersonal Relationships with Students	12	.36 - .66	33.29	13.37
III. Content Competency	12	.36 - .62	24.92	10.00

"Instructional Methods" had the greater number of items and accounted for the greater percent of common and total variance of the three factors.

The alignment of the items on the respective scales is presented in Table 7.

TABLE 7

The Alignment of Items on Each of the Four Subscale of the Teacher/Course Evaluation Instrument

	Subscales 1	2	3	4		Subscales 1	2	3	4
1.	*		*		24.				
2.			*		25.			*	
3.			*		26.		*		
4.	*		*		27.		*		
5.	*				28.				
6.	*				29.		*		
7.		*	*		30.	*		*	
8.		*			31.	*			
9.			*		32.		*		
10.					33.	*			
11.		*			34.	*			
12.	*				35.				*
13.	*				36.				*
14.		*			37.				*
15.	*				38.				*
16.	*		*		39.				*
17.			*		40.				*
18.	*				41.				*
19.			*		42.				*
20.		*			43.				*
21.			*		44.	*	*		
22.		*			45.	*	*		
23.									

Forty-one were included with items ten, twenty-three, twenty-four, and twenty-eight being excluded. Several of the items (numbers one, four, seven, sixteen, thirty, forty-four and forty-five) were included on two of the scales. Although the varimax rotation defines orthogonal factors reflected in the final factor scores (were they to be used), by scoring the scales with raw scores high inter-scale correlations were produced. This is a definite limitation in this study. However, from the standpoint of practical application of the findings (feedback to teachers) it was felt that raw scores rather than factor scores were more appropriate.

Table 8 is a presentation of the intercorrelations of the teacher/course evaluation subscales. All of the correlations were significant ($p < .05$) and rather high to be considered as measuring completely independent dimensions of teacher behavior or course attributes. Although the estimates of reliability (coefficient alpha) for

TABLE 8

Intercorrelations of Teacher/Course Evaluation Subscales (N=403) (Decimal Points Omitted)

Subscales	I	II	III	IV
I		71**	79**	60**
II			72**	56**
III				47**
IV				

** $p < .01$

each subscale were high, I - .89, II - .86, III - .86, IV - .77, the scales do reflect the problem of the intercorrelation of items found on more than one dimension in addition to the overlap of the dimensions themselves. Such data suggest that further refinement of the instrument is needed if these dimensions are to be made differentially effective one to another and meaningfully reflect student ratings of the various aspects.

Omnibus Personality Inventory - Form F. The following scales were included in the data analysis due to their relationships to the objectives of the study (i.e. differentiating college students according to values, attitudes, and opinions concerning the academic experience). The seven scales (their general descriptions taken from the test manual) included:

1. Thinking Introversion (TI). Persons scoring high on this measure are characterized by liking reflective thought and academic activities.
2. Theoretical Orientation (TO). High scores indicate preference for dealing with theoretical concerns and problems and use of the scientific method of thinking.
3. Estheticism (ES). High scores indicate diverse interests in artistic matters and activities, a high degree of sensitivity and response to esthetic stimulation.

4. Complexity (CO). High scores reflect an experimental and flexible orientation rather than fixed way of viewing and organizing phenomena.
5. Autonomy (AU). High scores express a tendency to be independent of authority as traditionally imposed through social institutions.
6. Religious Orientation (RO). High scores reflect a skepticism concerning religious belief and practices.
7. Anxiety Level (AL). High scores indicate a denial of feeling or symptoms of anxiety and do not admit being worried or nervous.
8. Intellectual Disposition Category (IDC). This is a composite score made up on the basis of 6 subscales scores which identify both the type and extent of commitment to general learning and intellectual activity. Eight classes are used with 1 and 2 indicative of broad intrinsic interests in intellectual or academic pursuits and 7 and 8 indicative of a limited and restricted orientation toward learning.

Data Collection Procedures

During the first week of Fall Quarter, students of teachers participating were asked to take part in the study and provide such information as sex, college membership and "expected grade" for the course. The distribution, completion, and collection of the forms were the responsibility of the teachers.

During the seventh and eigth weeks of the quarter the OPI was administered to each of the thirty sections on an individual basis during the regularly scheduled class period. The last week of the quarter prior to final examinations the teacher/course evaluation-instrument was given. Arrangements had been made in advance with the teachers to have approximately twenty minutes of one class period for this testing with teachers absent.

Additional student data were made available in several ways. English 160 grades were secured by having each teacher forward, at the end of the quarter, a copy of the assigned grades for each student in the participating sections. Fall grade

point average and American College Test (ACT) score information were released with the approval of the Director, Management Information Systems, for all students who had submitted correct student identification numbers. Of the possible 549 students in the study, 350 completed the form correctly enabling this information to be included in the data analysis.

Limitations of the Study

There were certain limitations of this study which should be noted.

1. Academic performance was indicated by the grade point average at the end of Fall Quarter. These marks were assigned by different teachers and although the shortcomings of such marks were realized, they were accepted as a relatively valid measure of student performance.
2. The American College Test scores were determined from one month to one year in advance of Fall Quarter entrance for the student. It was assumed that the relative position of subjects' scores one to another would not have to be significantly different had one mass testing been done for the entire sample.
3. The decrease in the sample size (from a potential 750 to 549) limits the generalizability of the findings.
4. Since the Omnibus Personality Inventory is a paper and pencil personality test, the difficulties inherent in this type of instrument were realized.
5. The high intercorrelation of the subscales (mean intercorrelation = .65) made it difficult to assume that they were measuring completely independent dimensions of teacher behavior. However, by squaring the mean intercorrelation value of .65, the coefficient of determination was found to be .42. This means that approximately 58% of the variance was unaccounted for thus suggesting that each of the scales was measuring relatively unique aspects of teacher behavior.

6. Since only 15 of the possible 180 teachers volunteered to participate in the study, such self-selection makes generalization of the findings to all sections of English 160 difficult if not impossible. However, demographically these 15 teachers were not significantly different from the total group.

Data Analysis

Omnibus Personality Inventory Subscale Scores. The findings presented in Table 9 show that students' teacher/course evaluations were independent of personality dimensions as measured by the OPI subscales. Although seven of the thirty-two correlations were significant ($p<.01$ and $p<.05$) in absolute values, the range was .105 to .229 indicating that the relationships found were low.

The lack of correlation between the OPI scores and the teacher/course evaluation subscales may be due to the sample used in this study. For example, in the studies reported by Yonge and Sassenrath (1968) and Weinstein and Bramble (1970), consisted exclusively of upperclass education majors enrolled in educational psychology courses, while this sample of students represented a wide variety of colleges and academic majors. The differences between these student groups may have contributed to potentially greater variability within classes and may have attenuated these relationships between students' personality scores and their evaluations.

Another explanation for this lack of significant correlations concerns the use and nature of the criterion instrument. Although no statistics concerning the validity or reliability of the instruments used elsewhere were available, possibly the evaluation instruments employed in the other studies were more capable of measuring the range of students' attitudes and opinions than demonstrated by the instrument used in the present study.

TABLE 9

Correlations of Teacher/Course Evaluation Subscale Scores with Selected Omnibus Personality Inventory Scores, American College Test Scores, Grade Point Average and the Variable Sex Across All Teachers (Decimal Points Omitted)

Variable	Teacher/Course Evaluation Subscales I	II	III	IV
OPI Subscales				
TI	229**	188**	179**	068
TO	066	006	006	-012
ES	152**	105*	078	016
CO	011	-048	-052	058
AU	007	-006	-032	013
RO	-041	-107*	-088	-032
AL	016	043	041	-068
IDC	-130**	-063	-076	-059

N = 403
*p< .05
**p< .01

American College Test Scores				
English	070	024	029	061
Mathematics	029	-053	-019	-010
Social Studies	024	034	-008	055
Natural Sciences	-010	-024	-013	004
Composite	034	-003	-002	031
Grade Point Average	114*	125*	121*	117*

N = 350
*p< .05

Sex	-002	120*	047	124*

* p< .05

American College Test Score. Of the twenty ACT correlations, none were significant, indicating that students' teacher/course evaluations were independent of ability as measured by ACT scores.

Although the instrumentation problem as presented in the discussion of the OPI findings is applicable here, the absence of significant correlations between the ACT scales and the components of the teacher/course evaluation was unexpected considering the findings in regard to grade point average and "actual grades" received (to be discussed later). Since English 160 was a required course, a wide range of abilities was expected in the sample of students. In addition, since ability is emphasized in the college environment, relationships between this variable and evaluations were hypothesized. This finding of independence between abiltiy measures and student ratings if replicated in other studies will encourage the use of such rating instruments.

Grade Point Average (GPA). As shown in Table 9, GPA was significantly positively correlated with all subscales. This finding is consistent with those reported by Weinstein and Bramble (1970) wherein they found students with higher grade point averages rated the instructor significantly higher than did students with low grade point averages. Unfortunately, their evaluation scale reflected an "omnibus" teacher rating and was not constructed with scales measuring various aspects of teacher behavior. Although these correlations were significantly correlated ($p < .05$) with subscales, in terms of absolute values the range of .114 to .125 was very low.

It should be noted that this sample consisted only of first quarter freshman students. As a result, grades in English 160 generally comprised from one-fourth to one-seventh of their total GPA. These findings should be considered tentative until further research is undertaken with students who have accumulated greater numbers of credit hours than represented by the first quarter freshmen in this sample.

Sex. The findings concerning the relationship of this variable to the various evaluative components are also presented in Table 9. Significant ($p < .05$) correlations were found for "Interpersonal Relationships with Students" and "Teacher Personality;" however, in terms of absolute values, they were not large (.120 and .124). Females, on these scales, gave their teachers higher mean evaluation than did males. Since the teacher sample had eight males and seven females these results cannot be attributed to a greater frequency of male teacher--female student interaction. An alternative possibility may be that females were more liberal in their ratings on these dimensions because of "knowing" their teachers better than males. On the basis of several conversations with a number of English 160 teachers, it was noted that females generally out-performed males on tests and overall were more interested and active in the course than were males. Another possible explanation might be that they felt more at ease being critical of the teacher's methods of instruction or competency than his or her "personal qualities."

Expected Grade. A significant F value ($p < .01$) was found only for "Interpersonal Relationships with Students" (Table 10). Using Scheffe's method of multiple comparisons of means (Table 11) it was found that students who expected A or B grades had significantly higher ($p \leq .05$) mean ratings than did students who expected

TABLE 10

Analysis of Variance of "Interpersonal Relationships With Students" (Subscale II) According to "Expected Grade"[a] in the Course Across All Teachers

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	931.48	2	465.74	7.25
Within Groups	29481.95	459	64.23	
Total	30413.45	461		

[a]Students who expected A totaled 47, B-270, C-144.

C grades. The differences between the A - B expected grade group when compared to the C group is difficult to explain since no similar differences were found on

TABLE 11

Differences Between Means of Groups Defined by "Expected Grade" in the Course on "Interpersonal Relationships With Students" (Subscale II) Using Scheffe's Method of Multiple Comparisons

	Expected Grade A $\bar{X}$ = 46.38 N = 47	Expected Grade B $\bar{X}$ = 45.03 N = 270	Expected Grade C $\bar{X}$ = 42.44 N = 144
A		1.35	3.94*
B			2.59*
C			
D			

* $p < .05$

the other subscales (since only four students received D grades they were deleted from this analysis). Further research, using instruments without high inter-correlations among the subscales and different samples of students should be considered before this variable is set aside.

Actual Grade. The data for this analysis are presented in Tables 12, 14, 16 and 18 in Appendix B. Significant F values ($p \leq .01$) were recorded for all subscales. Tables 13, 15, 17 and 19 (also Appendix B) are presentations of the Scheffe analyses which were calculated to compare the means of the groups. On all of the subscales, students who received A and B grades were significantly higher ($p \leq .05$) in the evaluations of the teacher than those students who received C and D grades. Although the subscales of this instrument were highly correlated, the relationship of "actual grades" received by students was significantly related to their evaluations of the teacher in the course as indicated by these findings. These results are in direct opposition to the studies reported by Remmers (1928) and Blum (1936) who reported no relationship between these variables.

These findings concerning "actual grades" should be considered within the context of the course under investigation. In English 160, students received

considerable feedback concerning the quality of their work. However, the grading practices were, in general, subjective rather than objective, and the teacher probably was more often seen as being more personally responsible for the grades distributed. Under these circumstances, students may have felt a stronger personal reaction (either positively or negatively) to the teacher than found with students in courses wherein more objective methods of evaluation and grade distribution were used.

College Membership. The findings presented in Table 20 show that students' evaluations were not independent of the variable "college membership." Significant differiences between the mean scores (Table 21) found between the College of Nursing and Fine and Professional Arts, Arts and Sciences, Business Management and Education with the mean for the students in Nursing being significantly higher than all on Subscale II - Interpersonal Relationships with Students. The mean for the students in the College of Education was significantly higher than that of the College of Business.

TABLE 20

Analysis of Variance of "Interpersonal Relationships With Students" (Subscale II) According to "College Membership"[a] of the Raters Across All Teachers

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Colleges	3466.19	5	693.24	10.19**
Within Colleges	32803.78	482	68.06	
Total	36269.97	487		

** $p < .01$

[a]College membership of the student raters: Arts and Sciences, 132; Fine and Professional Arts, 108; Business, 59; Education, 155; Nursing, 28; Health, Physical Education, 6.

TABLE 21

Difference Between Means of Groups Defined by "College Membership" on "Interpersonal Relationships with Students" (Subscale II) Using Scheffe's Method of Multiple Comparisons

	A&S	F&PA	BUS.	ED.	Nursing	HPER
	$\bar{X}$=44.72 N=132	$\bar{X}$=43.55 N=108	$\bar{X}$=41.71 N=59	$\bar{X}$=45.90 N=155	$\bar{X}$=54.36 N=28	$\bar{X}$=43.50 N=6
A&S		1.17	3.01	-1.18	-9.64*	1.22
F&PA			1.84	-2.35	-10.81*	0.05
BUS.				-4.19*	-12.65*	-1.79
ED.					-8.46*	2.40
NURSING						10.86
HPER						

* $p < .05$

Speculation about this finding is very limited; however, a possible explanation could be that students enrolled in the Colleges of Nursing and, to a small extent, Education look for and encourage interpersonal relationships with their instructors as part of their goals for the course. Thus, they feel some personal commitment to the teacher and their evaluations validly reflect the reality of their experiences. Students from the other colleges may possibly be more content-oriented and simply did not take the time or have the inclination to develop any personal relationships.

"Expected Versus Actual Grade" Differential. In this analysis, the "expected grade" in the course was compared with the "actual grade" received for all students. Three groups were formed, Higher--those students whose actual grade was higher than anticipated; Equal--those students whose expected grade was equal to that received; and Lower--those students whose grade was less than expected. One-way analysis of variance was used to compare the groups for each of the four subscales. As shown in Tables 22, 24, 26 and 28, (Appendix C) all tests were significant ($p \leq .01$). Scheffe's tests of multiple comparisons of means were calculated (Tables 23, 25, 27 and 28) and, in general, the same pattern of significant differences ($p \leq .05$)

emerged. Students who received grades higher than expected had significantly higher evaluations than did students in either of the other groups on Subscales "Instructional Methods," "Interpersonal Relationships with Students," the significant differences between the means occurred between the Higher and Lower groups and between the Equal and Lower groups. These data support the previously reported finding concerning "actual grade" received in the course wherein it was shown that grades were related to students' evaluations of the teacher or course. Once again, the problem of the high inter-correlations among the subscales confused the findings to the extent that speculation about which of the dimensions of teacher behavior was more significantly related to grades and grading practices was impossible.

The "expected grade" data were collected at the beginning of Fall Quarter. Thus, students were aware of the dichotomy (higher or lower) between this grade and the grades they were receiving regularly over the quarter. This finding suggests the possibility of an emotional reaction (positive feelings when higher and negative when lower) by students when their anticipated grades did not coincide with those they received. Further research concerning the relationship of "grade expectations" to teacher/course evaluations is needed in upperclass course offerings and content areas such as history, science classes and education courses to name a few.

Additional Data Analyses

On the basis of such variables as GPA, actual grades, and sex, inter-teacher comparisons were seen as appropriate. Since none of the commonly collected teacher variables had been secured in this study (i.e. instructional methods, interaction patterns with students, etc.) teachers were grouped using data which were available--the students' teacher/course evaluations. Mean values had been calculated for all subscales and teachers were divided into three groups labeled high (1), medium (2), and low (3) on the basis of these scores. A summing of

these values was done across scales for each teacher with the possible range being four to twelve. Six teachers had summed scores between twelve and ten, six were in the range from eight to six, and three scored a perfect four. These groups were labeled low (Group 3), medium (Group 2), and high (Group 1) respectively. The relationship of OPI subscale scores, ACT scores, grade point average, college membership, "expected grade" and actual grade" in the course, "expected versus actual grade" differential and sex to the components of the teacher/ course evaluation instrument were then calculated for each group.

The findings were that overall OPI and ACT scores, grade point average, college membership, "expected grade" in the course, "expected versus actual grade" differential and sex were uncorrelated with any of the evaluative subscales for the three teacher groups. In some few instances, significant statistics were found, however, there was a complete lack of discernible patterns overall (tables not included).

The findings for the variable "actual grade" in the course were different and rather interesting. The distribution of letter grades in each of the three groups is important. Teachers in Group I gave fourteen A's, fifty-two B's and Forty C or D grades. Group II teachers gave twenty-eight A's, seventy-six B's, and seventy-five C's or D's, while teachers in Group III gave eight A's, thirty-seven B's and one hundred thirty-seven C's or D's. Teachers in Groups I and II distributed their grades much more evenly than did those in Group III wherein nearly seventy-five percent of the students had C or lower grades.

When comparisons of the mean ratings by letter grade were calculated for teachers in Group I significant F values were found on all subscales (see Tables 30, 32, 34 and 36 in Appendix D). The results of the Scheffe tests showed that evaluations by students receiving A and B grades were significantly higher than students receiving C and D grades on Subscales "Interpersonal Relationships with Students" and "Teacher Personality" (Tables 33 and 37). Significant differences between the means of groups receiving B and C-D grades were found on Subscales

"Instructional Methods" and Content Competency" (Tables 31 and 33).

This relationship between "actual grades" and students teacher/course evaluations occurred only once in teacher Groups II and III. In Group II, students who received A grades had significantly higher mean ratings than did students who received C and D grades on "Teacher Personality." Significantly higher mean ratings existed between students who received B grades rather than C or D grades on "Interpersonal Relationships with Students" for teachers in Group III.

The significance of these findings concerning grades is that once teachers were grouped according to mean ratings across all dimensions, actual grades received in the course made no difference in overall evaluations except for teachers in Group I. For teachers in Groups II and III, however, grades received by the students did not have significant relationships to their ratings (except for two of the eight subscales). Thus, students of teachers in Group I used different criteria upon which to evaluate their instructors than did students of teachers in Groups II and III. In the latter groups, students classified these teachers as average or below average (summing across the four subscales) possibly in part because of their grading practices while teachers in Group I were seen as above-average teachers regardless of their grading habits. These findings suggest another variable influencing relationships. That is, if teachers are seen as "good" or "above-average" overall by their students, then the criteria used by students for a total evaluation are not related to the actual grade they received in the course (since significant mean differences by letter grade were evident in the student ratings yet these teachers had the highest mean ratings in the sample of teachers). In comparison, when students saw a teacher as "average" or "below average" they may have used as their criteria grading practices or merely reflected a discontent with the teacher's attitude as demonstrated in his grading pattern. Further research is necessary to determine what specific or general criteria students in this course use to rate teachers. If grades do not have that significant relationship to "good teachers" performance, then what variables are the students

using?

Discussion

As indicated previously, one of the limitations of this study and others like it is the relatively high intercorrelations of thesubscales of the teacher/course evaluation. A possible method for decreasing this high intercorrelation may be to have students, through separate instructions, evaluate portions of teacher behavior individually. This would allow them to focus their thoughts on homogeneous scales reflecting teachers' behavior or aspects of the course without the problem of continually changing their mental set as occurs when heterogeneous grouping of items is done. This procedure might act to lower the intercorrelations among the scales while increasing the stability of the ratings over time.

When teacher/course evaluation instruments have been revised, questions concerning reliability and not validity have been most often given attention. Rarely, if at all, have questions concerning the basic validity of the instruments been asked. Two types of validity should be distinguished at this point. In one type, validity is concerned with whether students' ratings honestly reflect the teacher behavior being evaluated. In the second type, validity is concerned with whether the behavior the students are rating in a given item is appropriate to the content area under investigation (content validity). In other words, should teachers in English be rated on the same scales measuring "methods of instruction" as are teachers of physical education? Should teachers be measured on "interpersonal relationships with students" when such abilities are not valued in all subject matter areas? These questions should be considered as integral to proper instrument development, revision and use.

One approach to the question of validity would be to use a two stage approach to teacher/course evaluation instrument construction. In the first stage, inter-departmental or inter-college differences between courses and teachers could be defined through the use of the "behavioral approach" (Solomon, Issaacson). This

data could then be used to determine which teacher behaviors or course aspects are unique to specific content areas. In the second part, the "evaluative approach" could be used to develop items appropriate to the appraisal of behaviors in each area. This practice would be a first step toward establishing content validity for such rating instruments.

Considering the number of variables which were significantly related to the teacher/course evaluations in this study, the use of such instruments for the purpose of faculty promotions, tenure, and pay raises might be cautiously approached. Eventually, for example, it might be better if evaluations were analyzed and reported separately for each assigned grade student group to appraise the overall effectiveness of the teacher. However, much research (instrument development, other samples, etc.) is needed before such corrective measures are undertaken.

The possible classroom application of these findings concerning student variable correlates of teacher/course evaluations should not be overlooked. Although personality and ability measures were uncorrelated in this sample of teachers, this does not negate the possibility of such correlations with specific teachers as suggested by the findings of Yonge and Sassenrath (1968). In addition, the other variables included in this investigation such as grade point average, sex, and "expected grade" in the course might be of value to individual teachers for their use in the planning of educational strategies. If, for example, a teacher found that consistently lower evaluations of his "instructional methods" were provided by students with high grade point averages, this would suggest that he should alter his methods to better accomodate the needs of these students or request that they be placed in classes with teachers who offer more effective techniques. This approach to the use of feedback from teacher/course evaluations focuses upon the possible grouping of students with teachers who are known to possess certain qualitites or characteristics desired by particular students. Since little research has been done in this area to show that achievement gains are effected by such clustering, efforts should be undertaken to further clarify the relationship between teacher-

student matching and student achievement. Of particular importance is the question of whether teacher/course evaluation ratings can be used effectively as a means for grouping students with teachers.

Further research should be undertaken to determine what criterion variables students use to evaluate "above-average" teachers since in this study these teachers were placed in this category by students who used criteria other than grades received in the course or teacher grading practices. What teacher characteristics were the students using for such classification? Since teachers who received average and below-average ratings were possibly classified to a large degree on the basis of these variables, why were above-average exempt from this? Such a finding suggests that a "master teacher" group may exist as determined by student ratings and further research should be done to determine their characteristic patterns of behavior both in and out of the classroom environment.

APPENDIX A

TEACHER/COURSE EVALUATION FORM

This teacher --	STRONGLY DISAGREE	DISAGREE	NEUTRAL	AGREE	STRONGLY AGREE
1. - was able to effectively relate the course materials to the broader field of knowledge.	1	2	3	4	5
2. - expressed himself (herself) clearly when writing.	1	2	3	4	5
3. - showed a good working knowledge of the rules of grammar.	1	2	3	4	5
4. - made valid interpretations of the readings (e.g. poems, novels).	1	2	3	4	5
5. - capably related information from other fields to the course material.	1	2	3	4	5
6. - presented the material in an interesting manner.	1	2	3	4	5
7. - was able to effectively communicate the rules of good writing.	1	2	3	4	5
8. - presented the subject matter in too complex a manner.	1	2	3	4	5
9. - was able to effectively synthesize, integrate, and summarize the subject matter.	1	2	3	4	5
10. - covered the material too slowly.	1	2	3	4	5
11. - by his (her) actions, seemed to view teaching as a chore or routine activity.	1	2	3	4	5
12. - showed an engaging enthusiasm for the subject.	1	2	3	4	5

TEACHER/COURSE EVALUATION FORM

	STRONGLY DISAGREE	DISAGREE	NEUTRAL	AGREE	STRONGLY AGREE
13. - was concerned about stimulating students' curiosity in the subject.	1	2	3	4	5
14. - often used class time with discussion of irrelevent or meaningless topics.	1	2	3	4	5
15. - used a style of lecturing which was dull and uninteresting.	1	2	3	4	5
16. - effectively used a variety of instructional methods which were appropriate to the course material.	1	2	3	4	5
17. - listened attentively to students' questions, comments, and remarks during class.	1	2	3	4	5
18. - effectively used mimicry, anedotes, and/or a general herminess to enliven the class period.	1	2	3	4	5
19. - was able to communicate clearly the directions for individual assignments.	1	2	3	4	5
20. - relied too heavily on student performance (e.g. talking, answering questions, etc.) in class as the primary basis of grading.	1	2	3	4	5
21. - was generally well prepared for class.	1	2	3	4	5
22. - was too inflexible concerning his (her) right to control the in-class discussions and activities.	1	2	3	4	5
23. - should have relied more heavily on objective tests for grading purposes.	1	2	3	4	5

TEACHER/COURSE EVALUATION FORM

	STRONGLY DISAGREE	DISAGREE	NEUTRAL	AGREE	STRONGLY AGREE
24. - was punctual about meeting class.	1	2	3	4	5
25. - was available for students to talk to when not in class.	1	2	3	4	5
26. - too often forced his (her) ideas or opinions on the class.	1	2	3	4	5
27. - was sometimes unfair in the grading of students' work.	1	2	3	4	5
28. - was punctual about dismissing class.	1	2	3	4	5
29. - was threatening and caused students to be afriad of speaking in class.	1	2	3	4	5
30. - presented the material so that it was intellectually challenging to the student.	1	2	3	4	5
31. - was a monotonous and dull speaker.	1	2	3	4	5
32. - often made individual students feel uncomfortable or embarrassed in class.	1	2	3	4	5
33. - was able to stimulate interesting class discussions.	1	2	3	4	5
34. - displayed only a test related knowledge.	1	2	3	4	5

EACH OF THE FOLLOWING SCALES HAVE FIVE NUMBERS ON THEM WITH A DESCRIPTIVE ADJECTIVE OR PHRASE ON EACH SIDE. YOU ARE TO DECIDE WHICH ADJECTIVE OR PHRASE BEST DESCRIBES THE TEACHER IN THIS COURSE AND THEN HOW STRONGLY YOU WOULD APPLY THE DESCRIPTION TO HIM (HER). YOU ARE TO SELECT THE APPROPRIATE NUMBER WHEN A 3 INDICATES "UNCERTAIN"; 4 AND 5 INDICATE INCREASING DEGREES OF AGREEMENT WITH THE ADJECTIVE OR PHRASE ON THE RIGHT; AND 2 OR 1 INDICATE INCREASING DEGREES OF AGREEMENT WITH THE ADJECTIVE OR PHRASE ON THE LEFT.

35. uneasy and hesitant...1 2 3 4 5.....calm and relaxed

36. lacked sense of humor.1 2 3 4 5.displayed sense of humor

37. sincere.....1 2 3 4 5.....insincere

38. self-confident.....1 2 3 4 5.....uncertain

39. cold and impersonal...1 2 3 4 5.....warm and friendly

40. flexible.....1 2 3 4 5.....rigid

41. threatening.....1 2 3 4 5.....non-threatening

42. formal.....1 2 3 4 5.....informal

43. unsociable.....1 2 3 4 5.....sociable

44. How much do you feel you have learned from this teacher?

1. very little
2. a small amount
3. a fair amount
4. quite a bit
5. a great amount

45. How much do you think you would like the instructor in this course as a personal friend?

1. not at all
2. slightly
3. somewhat
4. quite a bit
5. very much

TABLE 12

Analysis of Variance of "Instructional Methods" (Subscale I) According to "Actual Grade"[a] Received in the Course Across All Teachers

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	2736.12	3	912.04	9.67**
Within Groups	43688.83	463	94.36	
Total	46424.95	466		

** p< .01

[a]Number of students who received A totaled 50, B-165, C-220, and D-32.

TABLE 13

Differences Between Means of Groups Defined by "Actual Grade" Received in the Course on "Instructional Methods" (Subscale I) Using Scheffe's Method of Multiple Comparisons

	Actual Grade A $\bar{X}$ = 57.00 N = 50	Actual Grade B $\bar{X}$ = 56.32 N = 165	Actual Grade C $\bar{X}$ = 52.26 N = 220	Actual Grade D $\bar{X}$ = 49.22 N = 32
A		0.68	4.73*	7.78*
B			4.06*	7.10*
C				3.04
D				

* p< .05

TABLE 14

Analysis of Variance of "Interpersonal Relationships With Students' (Subscale II) According to "Actual Grade"[a] Received in the Course Across All Teachers

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	3791.76	3	1263.92	21.99**
Within Groups	26600.36	463	57.45	
Total	30392.12	466		

** p< .01

[a]Number of students who received A totaled 50, B-165, C-220, and D-32.

TABLE 15

Differences Between Means of Groups Defined by "Actual Grade" Received in the Course on "Interpersonal Relationships With Students" (Subscale II) Using Scheffe's Method of Multiple Comparisons

	Actual Grade A $\bar{X}$ = 48:50 N = 50	Actual Grade B $\bar{X}$ = 46.95 N = 165	Actual Grade C $\bar{X}$ = 42.36 N = 220	Actual Grade D $\bar{X}$ = 38.91 N =32
A		1.55	6.14*	9.59*
B			4.59*	8.04*
C				3.45
D				

* p < .05

TABLE 16

Analysis of Variance of "Content Competency" (Subscale III) According to "Actual Grade" Received in the Course Across All Teachers

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	1298.64	3	432.88	9.65**
Within Groups	20770.13	463	44.86	
Total	22068.77	466		

** p< .01

[a]Number of students who received A totaled 50, B-165, C-220, and D-32.

TABLE 17

Differences Between Means of Groups Defined by "Actual Grade" Received in the Course on "Content Competency" (Subscale III) Using Scheffe's Method of Multiple Comparisons

	Actual Grade A $\bar{X}$ = 47.48 N = 50	Actual Grade B $\bar{X}$ = 47.72 N = 165	Actual Grade C $\bar{X}$ = 44.75 N = 220	Actual Grade D $\bar{X}$ = 42.66 N = 32
A		-0.38	2.73	4.82*
B			2.97*	5.06*
C				2.09
D				

* p < .05

TABLE 18

Analysis of Variance of "Teacher Personality" (Subscale IV) According to "Actual Grade"[a] Received in the Course Across All Teachers

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	1690.72	3	563.57	17.95**
Within Groups	14533.54	463	31.39	
Total	16224.26	466		

** p < .01

[a]Number of students who received A totaled 50, B-165, C-220, and D-32.

TABLE 19

Differences Between Means of Groups Defined by "Actual Grade" Received in the Course on "Teacher Personality" (Subscale IV) Using Scheffe's Method of Multiple Comparisons

	Actual Grade A $\bar{X}$ = 38.70 N = 50	Actual Grade B $\bar{X}$ = 37.33 N = 165	Actual Grade C $\bar{X}$ = 34.60 N = 220	Actual Grade D $\bar{X}$ = 31.59 N = 32
A		1.37	4.10*	7.11*
B			2.73*	5.74*
C				3.01*
D				

* p < .05

TABLE 22

Analysis of Variance of "Instructional Methods" (Subscale I) According to "Actual Versus Expected Grade"[a] Differential (Higher, Equal, and Lower) Across All Teachers

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	1049.49	2	524.75	5.36**
Within Groups	45415.45	464	97.88	
Total	46464.94	466		

** $p < .01$

[a]Student membership by group: Higher, 77; Equal, 213; and Lower, 177.

TABLE 23

Differences Between Means of Groups Defined According to "Actual Versus Expected Grade" Differential on "Instructional Methods" (Subscale I) Using Scheffe's Method of Multiple Comparisons

	Higher	Equal	Lower
	$\bar{X}$ = 57.21 N = 77	$\bar{X}$ = 53.65 N = 213	$\bar{X}$ = 52.85 N = 177
Higher		3.56*	4.36*
Same			0.80
Lower			

* $p < .05$

TABLE 24

Analysis of Variance of "Interpersonal Relationships With Students" (Subscale II) According to "Actual Versus Expected Grade"[a] Differential (Higher, Equal, and Lower) Across All Teachers

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	1421.41	2	710.70	11.24**
Within Groups	29345.87	464	63.25	
Total	30767.28	466		

** $p < .01$

[a]Student Membership by group: Higher, 77; Equal, 213; and Lower, 177.

TABLE 25

Differences Between Means of Groups Defined According to "Actual Versus Expected Grade Differential" on "Interpersonal Relationships With Students" (Subscale II) Using Scheffe's Method of Multiple Comparisons

	Higher	Equal	Lower
	$\bar{X}$ = 47.90 N = 77	$\bar{X}$ = 44.41 N = 213	$\bar{X}$ = 42.75 N = 177
Higher		3.49*	5.15*
Equal			1.66
Lower			

* $p < .05$

TABLE 26

Analysis of Variance of "Content Competency" (Subscale III) According to "Actual Versus Expected Grade"[a] Differential (Higher, Equal, Lower)Across All Teachers

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	604.05	2	302.02	6.52**
Within Groups	21493.42	464	46.32	
Total	22097.47	466		

** $p < .01$

[a]Student membership by group: Higher, 77; Equal, 213; and Lower, 177.

TABLE 27

Differences Between Means of Groups Defined According to "Actual Versus Expected Grade Differential" on "Content Competency" (Subscale III) Using Scheffe's Method of Multiple Comparisons

	Higher	Equal	Lower
	X = 48.31 N = 77	X = 45.81 N = 312	X = 44.97 N = 177
Higher		2.50*	3.34*
Equal			0.84
Lower			

* $p < .05$

TABLE 28

Analysis of Variance of "Teacher Personality" (Subscale IV) According to "Actual Versus Expected Grade"[a] Differential (Higher, Equal, Lower) Across All Teachers

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	646.53	2	323.26	9.68**
Within Groups	15498.55	464	33.40	
Total	16145.08	466		

** $p < .01$

[a]Student Membership by group: Higher, 77; Equal, 213; and Lower, 177.

TABLE 29

Differences Between Means of Groups Defined According to "Actual Versus Expected Grade Differential" On "Teacher Personality" (Subscale IV) Using Scheffe's Method of Multiple Comparisons

	Higher	Equal	Lower
	$\bar{X}$ = 37.94 N = 77	$\bar{X}$ = 36.07 N = 213	$\bar{X}$ = 34.54 N = 177
Higher		1.87	3.40*
Equal			.153*
Lower			

* $p < .05$

TABLE 30

Analysis of Variance of "Instructional Methods" (Subscale I) According to "Actual Grade" Received in the Course Across All Teachers with High Mean Evaluation Ratings (Group I)[a]

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	548.94	2	274.47	5.32**
Within Groups	5315.30	103	51.60	
Total	5864.24	105		

** $p < .01$

[a]Student raters receiving A, 14; B, 52; C or D, 40.

TABLE 31

Differences Between Means of Groups Defined by "Actual Grade" Received in the Course on "Instructional Methods" (Subscale I) Across Teachers with High Mean Evaluation Ratings (Group I) Using Scheffe's Method of Multiple Comparisons

	Actual Grade A $\bar{X}$ = 60.50 N = 14	Actual Grade B $\bar{X}$ = 59.44 N = 52	Actual Grade C and D $\bar{X}$ = 55.03 N = 40
A		0.06	5.47
B			4.41*
C-D			

* $p < .05$

TABLE 32

Analysis of Variance of "Interpersonal Relationships With Students" (Subscale II) According to 'Actual Grade' Received in the Course Across All Teachers With High Mean Evaluation Ratings (Group I)[a]

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	476.23	2	238.12	8.85**
Within Groups	2771.43	103	26.91	
Total	3247.66	105		

** $p < .01$

[a]Student raters receiving A, 14; B, 52; C or D, 40.

TABLE 33

Differences Between Means of Groups Defined by "Actual Grade" Received in the Course on "Interpersonal Relationships With Students" (Subscale II) Across Teachers with High Mean Evaluation Ratings (Group I) Using Scheffe's Method of Multiple Comparisons

	Actual Grade A $\bar{X}$ = 51.79 N = 14	Actual Grade B $\bar{X}$ = 50.21 N = 72	Actual Grade C and D $\bar{X}$ = 46.30 N = 40
A		1.58	5.49*
B			3.91*
C-D			

* $p < .05$

TABLE 34

Analysis of Variance of "Content Competency" (Subscale III) According to "Actual Grade" Received in the Course Across All Teachers with High Mean Evaluation Ratings (Group I)[a]

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	230.22	2	115.11	4.11*
Within Groups	2884.66	103	28.00	
Total	3114.88	105		

* $p < .05$

[a]Student raters receiving A, 14; B, 52; C or D, 40.

TABLE 35

Differences Between Means of Groups Defined by "Actual Grade"Received in the Course on "Content Competency" (Subscale III) Across Teachers with High Mean Evaluation Ratings (Group I) Using Scheffe's Method of Multiple Comparisons

	Actual Grade A $\bar{X}$ = 50.14 N = 14	Actual Grade B $\bar{X}$ = 50.33 N = 52	Actual Grade C and D $\bar{X}$ = 47.25 N = 40
A		-0.19	2.89
B			3.08*
C-D			

* $p < .05$

TABLE 36

Analysis of Variance of "Teacher Personality" (Subscale IV) According to "Actual Grade" Received in the Course Across All Teachers with High Mean Evaluation Ratings (Group I)[a]

Source of Variation	Sums of Squares	Degrees of Freedom	Mean Square	F
Between Groups	322.94	2	161.47	6.07**
Within Groups	2740.83	103	26.61	
Total	3063.77	105		

** $p < .01$

[a]Student raters receiving A, 14; B,52; C or D, 40.

TABLE 37

Differences Between Means of Groups Defined by "Actual Grade" Received in the Course on "Teacher Personality" (Subscale IV) Across Teachers with High Mean Evaluation Ratings (Group I) Using Scheffe's Method of Multiple Comparison

	Actual Grade A $\bar{X}$ = 40.71 N = 14	Actual Grade B $\bar{X}$ = 39.81 N = 52	Actual Grade C and D $\bar{X}$ = 36.45 N = 40
A		00.90	3.86*
B			2.96*
C-D			

* $p < .05$

REFERENCES

Blum, M.L. An investigation of the relation existing between students' grades and their ratings of the instructors ability to teach. Journal of Educational Psychology, 1936, 27, 217-221.

Heist, P. and Yonge, G. The Omnibus Personality Inventory: form f, manual. New York: the Psychological Corporation, 1968.

Issacson, R.L., McKeach, W.J. and Milholland, J.E. Correlation of teacher personality variables and student ratings. Journal of Education Psychology, 1963, 54, 110-117.

Issacson, R.L. et al. Dimensions of student evaluations of teaching. Journal of Educational Psychology, 1964, 55, 344-351.

Rees, R.D. Dimensions of Students' Points of View in Rating College Teachers. Journal of Educational Psychology, 1969, 60, 476-482.

Remmers, H.H. Experimental data on the Purdue Rating Scale for Instruction. Educational Administration and Supervision, 1927, 13, 519-527.

Weaver, C.H. Instructor rating by college students. Journal of Education Psychology, 1960, 51, 21-25.

Weinstein, P. and Bramble, W.J. Student press: student course ratings as a function of student variables. Paper read at the 1971 American Educational Research Association Meeting, New York, February, 1971.

Yonge, G.D. and Sassenrath, J.M. Student personality correlates of teacher ratings. Journal of Educational Psychology, 1968, 59, 44-59.